Philosophy of probability

Wenmackers, S.

Chapter 4

Ultralarge and Infinite Lotteries

Our minds are finite, and yet even in these circumstances of finitude we are surrounded by possibilities that are infinite, and the purpose of human life is to grasp as much as we can out of the infinitude.
Alfred North Whitehead (Price, 1954, p.160)

'Can you do Addition?' the White Queen asked. 'What's one and one and one and one and one and one and one and one and one and one?'
'I don't know,' said Alice. 'I lost count.'
Lewis Carroll (1871, Chapter IX)

One, two, three ... infinity.
George Gamow (1947)

4.1 Introduction

This chapter is intended to summarize the findings of Chapters 2 and 3, to emphasize the analogy between the problems and the solutions discussed there, and to develop the epistemology of an infinite lottery.

The problem of describing a fair infinite lottery (subsection 4.1.1) and the Lottery Paradox (subsection 4.1.2) are usually considered to be unrelated issues, appearing in different subdisciplines of philosophy. In this chapter, however, we will emphasize the analogy between them. This approach allows us to diagnose them as suffering from an 'adding problem', which is in both cases due to the accumulation of rounding errors (section 4.2). Because the diagnosis is the same, we can proscribe the same treatment, too: we will achieve this with the help of infinitesimals, which are available in non-standard analysis (NSA) (section 4.3). For each of the two cases, we can select an approach to NSA that suits the application best. In section 4.4, we will combine our findings to analyze the epistemology of an infinite lottery.

4.1.1 Case 1: Failure of Countable Additivity in an infinite lottery

In standard probability theory, with the axioms as introduced by Kolmogorov (1933), it is not possible to describe a fair lottery on a countably infinite sample space.

To see this, consider $\mathbb{N}$ as the sample space (usually denoted by Ω). The axiom of Normalization requires that the probability of the full set of tickets equals one. To model a fair lottery means that all the tickets have the same probability. If this is so, the probability of a single ticket cannot be any number larger than zero, for then the infinite sum ranging over all the tickets diverges, and either Normalization or Countable Additivity (CA) fails. However, the probability cannot be zero either, because then the infinite sum of all the individual probability values evaluates to zero, which is not equal to one as demanded by Normalization and CA. Hence, we cannot assign *any* probability to the individual tickets: Kolmogorov's approach to probability theory cannot describe a fair infinite lottery.

It is puzzling that the formalism is not flexible enough to deal with this case, in particular since a countably infinite lottery corresponds to a sample space of the lowest infinite cardinality.

Various solutions have been proposed to solve this problem:

(1) Give up the property of Countable Additivity.

(2) Give up the axiom of Normalization.

(3) Deny that there exists a fair lottery on $\mathbb{N}$.

(4) Assign an infinitesimal probability to the single tickets, rather than zero.

The most prominent advocate of option (1) was de Finetti (1974), who argued to replace CA by the weaker requirement of Finite Additivity. Solution (2) was proposed by Rényi (1955). If we hold on to Kolmogorov's system, including Normalization and CA, we arrive at option (3): probability measures on countable domains are always biased. This solution has some strange consequences, to say the least (see for instance Kelly, 1996, p. 321–323).

We find solution (4), advocated by Lewis and Skyrms (1980), to be the most plausible one. It is not immediately clear which changes have to be made in Kolmogorov's axioms in order to realize this solution. A detailed analysis (as the one presented in Chapter 2) shows that not only has the range of the probability function to be adapted (from the standard unit interval of the real numbers to a non-standard set), but that even then the property of Countable Additivity cannot hold (although it can be replaced by another form of infinite additivity). Hence, it turns out that (1) and (4) are not completely independent solutions after all.

4.1.2 Case 2: Failure of Conjunction Principle for rational beliefs concerning a large, finite lottery

The probability assignment for a fair, finite N-ticket lottery with exactly one winner is trivial: every ticket has a probability of winning equal to $1/N$. From an epistemological viewpoint, however, the case where $N \gg 1$ is problematic: a paradox can be obtained for rational beliefs about such a lottery.

Consider a very large, but finite fair lottery. If you own just one ticket, it may seem rational for you to say: "My single ticket will not win." Suppose that you receive an additional ticket for the same lottery. It may still be rational for you to say: "My two tickets will not win." Likewise for three tickets, and so on. If you generalize the previous statement to "Likewise for all (N) tickets", you obtain a paradox, for you know the lottery to have a winner for certain.

This Lottery Paradox was first constructed by Kyburg (1961), who concluded that the Conjunction Principle (CP) does not hold for rational acceptability. Many later authors have tried to rescue at least part of CP (Wheeler, 2007).

Although we are in favor of a formal approach to epistemology, the popular threshold-based model for rational beliefs—which states that it is rational to believe a statement if the probability of that statement is larger than some positive threshold smaller than one—does not seem adequate. In particular, it does not do justice to the intrinsic vagueness of both the Lockean thesis—the thesis stating that it is rational to believe a statement if the probability of that statement is 'sufficiently close to unity'—and the Lottery Paradox itself, which, after all, only occurs in cases in which the number of tickets is 'sufficiently large'. The phrasings 'sufficiently close to unity' and 'sufficiently large' are strongly suggestive of a vague and contextual element, both of which are absent in threshold-based models.

Intuitively, we would expect a weakened version of CP to hold, allowing us to aggregate at least 'a few'—again a vague and context-dependent term—rational beliefs. Applied to the example of the Lottery Paradox, the maximal generalization that can rationally be obtained is: "Any set containing a few ticket will not win"—not 'many', in particular not 'all'. In the threshold-based model, however, CP fails and there is no weakened version of it available either.

4.2 The analogy between the two cases

Though appearing in different contexts, both cases have common characteristics:

(a) They both involve the assignment of fair probabilities to a denumerable lottery.

(b) They deal with either a 'sufficiently large' or an 'infinite' sample space, two concepts that are intimately related.

(c) They both involve a vague distinction.

(d) They both involve a limit process.

(e) They both pose an adding problem.

Whereas characteristic (a) is immediately clear, we will now provide further motivations for the statements (b)–(e).

Characteristic (b) claims that the properties of being finite but sufficiently large and of being infinitely large are intimately related. This is not on original claim: Lavine (1995) has argued that our whole concept of infinity has to have originated, somehow, from our experience with finite quantities (since our senses do not allow us to experience any infinite quantity). In particular, Lavine sees 'infinity' as a way of dealing with the 'indefinitely large' numbers. Although the former has the benefit of being a context-independent concept, it is not strictly necessary, and Lavine (1995) argues to practice mathematics only in terms of the latter, finite concept. Those who do not endorse a finitistic attitude towards mathematics may employ Lavine's claim for the analogy between 'indefinitely large' and 'infinite' in both directions: whereas a generalization occurs from the finite to the infinite realm in the development of NSA, we will apply NSA to the infinite lottery first and use this result as an inspiration for dealing with the Lottery Paradox, involving a large but only finite domain.

The presence of vague distinctions (c) has already been demonstrated for the Lottery Paradox, but may seem an unwarranted claim when considering the infinite lottery problem. Admittedly, the claim there is of a more subtle nature: it says that the distinction between finite and infinite is vague. As we will see, in NSA there are infinite hypernatural numbers, which are larger than any finite natural number. Just like there is no last (*i.e.* largest) finite natural number, there is no first (*i.e.* smallest) infinite hypernatural one. Hence, the distinction between finite and infinite, clear as it may seem, does not amount to a sharp threshold between the two, but a transition that can only be indicated by points of ellipsis.

An ellipsis often indicates a type of limit process, which brings us to (d). The limit process involved in the infinite lottery seems clear: we want to idealize a fair lottery on N as N becomes larger, in such a way that the answer is no longer dependent on how large N is precisely. Kolmogorov's framework—in particular, his Axiom of Continuity, which has CA as a consequence—implies using a classical limit, which is always defined on the real numbers. However, this is not the only limit operation mathematics has to offer. As we will see in the next section, one approach to NSA has a limit operation available that evaluates to an infinitesimal rather than zero, for the winning odds of a single ticket in an infinite lottery. Likewise, we may interpret another approach to NSA as blocking the generalization of the statements in the Lottery Paradox to a statement about 'many' (in particular, 'all') tickets, while allowing the corresponding statement in terms of 'a few' tickets.

The adding problem referred to in (e) refers to the failure of CA on the one hand and to the failure of CP, even in a weakened form, on the other. We will now show that the problem seems to be due to the accumulation of rounding errors in both cases. To see this, all we need is a basic understanding of error propagation. For instance, if two or more rounded values are added, the total error on the sum is equal to the sum of the errors on the terms. On its own, a rounding error it just that: one small error. However, in a sum with many terms, all of which have a small error associated with them, the total error is of the form 'many $\times$ small', which is not guaranteed to be small.

In the next section, we will investigate whether we can employ the analogies observed here to resolve the two issues with the same solution strategy.

4.3 Solution using non-standard analysis

The solution we will pursue is to represent the aforementioned rounding errors by infinitesimals. Whereas it has long been thought that infinitesimals are an inconsistent concept, Robinson (1966) has developed NSA as a framework that allows us to deal with infinitesimals—as well as infinite numbers—consistently.

4.3.1 Infinite additivity for a fair lottery on $\mathbb{N}$

Because the infinite lottery case seems the most natural one to apply NSA to, we will consider this first. In fact, there are multiple approaches to NSA that can be followed here. A model in terms of free ultrafilters can be constructed, as is done in Chapter 2. One may also start from the numerosity theory of Benci and Di Nasso (2003b). Since this approach starts from an axiomatic basis, it allows us to stress the concepts behind the theory rather than the technical details, as seems appropriate within the scope of a relatively short chapter.

Numerosity offers a way to assign sizes to sets such that a set is guaranteed to have a larger size than any of its proper subsets (Benci and Di Nasso, 2003b). This entails that sets of equal cardinality may have different numerosities, whereas the opposite does not hold. In particular, numerosities provide a way of distinguishing the size of infinite subsets within the natural numbers. In order to turn the numerosity function into a probability function, all that remains to be done is to take care of Normalization. Because numerosities are hyperreal numbers, they are closed under division and, thus, Normalization poses no difficulty. The numerosity of the full set of natural numbers is denoted by α (which is an infinite hypernatural number). Hence, we divide the numerosity of a subset of $\mathbb{N}$ by α to define a probability function that is defined on the whole power set of $\mathbb{N}$. Because the numerosity of a singleton is 1, the probability of a single ticket in a fair lottery on $\mathbb{N}$ is $1/\alpha$, an infinitesimal hyperrational number.[1]

What this approach amounts to in terms of a limit operation is the alpha-limit, which can be considered as a finer limit than the classical limit of standard analysis: sequences with equal classical limits, may have infinitesimally different alpha-limits. In particular, whereas the empty set and any finite subset of $\mathbb{N}$ both result in zero as the limiting probability in the classical approach (called asymptotic density, see *e.g.* Schurz and Leitgeb, 2008), only the finite set has probability zero in terms of the alpha-limit. Thus, for the infinite lottery, NSA makes it possible to assign infinitesimal probabilities to each ticket, which in turn allows us to have an infinite additivity property—though not CA (for details, see Chapter 2).

[1] A hyperrational number is a fraction of hypernatural numbers. The set of hyperrational numbers form a proper subset of the hyperreal numbers.

4.3.2 Conjunction Principle within a level for Stratified Beliefs concerning an ultralarge lottery

For the finite lottery, the application of NSA is slightly more indirect. After all, it is not the probability function itself that is unable to distinguish between the winning odds of an almost empty set of tickets and that of the empty set. The problem seems to arise only when we start to form beliefs based on those probability values. We propose a threshold-free model for beliefs based on an alternative formulation of NSA, called relative or stratified analysis (Hrbacek, 2007). We derive a rule that tells us when CP holds and when it is violated: it is allowed within a level, which is a concept from relative analysis intended to model the mesoscopic scale of magnitude. These levels always contain only a finite number of numbers, though some may be 'finer'—*i.e.* contain more numbers, both larger and smaller ones. A level always contains the number 1 and all quantities that are neither too small nor too large to be observable (depending on the context), but these quantities are not known with infinite precision. Real numbers that are larger than any number of a certain level are relatively infinite or ultralarge numbers compared to that level; positive real numbers that are smaller than any non-zero number belonging to the level are relative infinitesimals or ultrasmall numbers. In our model, the probability of winning for a single ticket in a fair, ultralarge lottery is a relative infinitesimal; this means that it is indistinguishable from zero (on that level). Details of the solution may be found in Chapter 3.

One may wonder whether it is really rational to believe that some non-empty set of possible outcomes will not occur, in exactly the same way as one believes that the impossible event will not occur. The answer is 'it depends', which is precisely the reason for the context-dependent element of levels in the model. If an agent is interested in, for instance, comparing the winning odds of someone owning a small set of tickets to that of someone who has no ticket at all, the very nature of the question is aimed at the difference between a small non-zero and a zero probability. In such a case, it is not likely that the agent will ignore the non-zero probability of a non-empty set. However, it may be rational for the agent to approximate an almost impossible event as having zero probability when comparing it to a certain or almost certain event. In that case, the error introduced by the approximation is of no significance to the conclusion.

The role of a level is similar to that of quantifier domain restriction, known from the philosophy of language.[2] Typically, when we refer to 'everbody' or 'everywhere', we do not mean 'all the people on Earth' or 'the whole Universe', but rather a more specific subgroup of people or a certain region on Earth, respectively. Clearly, the meaning of such a quantifier depends on the context in which the statement is uttered. Likewise, a level makes it possible to quantify only over the numbers that are observable in a given context.

[2]Thanks to Lorenz Demey for this suggestion. For a survey of quantifier domain restriction, consult for instance Stanley and Szabó (2000). The relation between domain restriction and vagueness is explored by Pagin (2010).

4.3.3 Visual summary of the solution

Figure 4.1 visualizes the analogy between the two cases. Figure 4.1(a) shows how the hyperrational probability values assigned to a fair lottery on $\mathbb{N}$ correspond to the real-valued approximation of (generalized) asymptotic density (as discussed by Schurz and Leitgeb, 2008). Whereas the former are fully additive, the latter is only finitely additive. Figure 4.1(b) shows why the full additivity of probabilities does not necessarily lead to full 'additivity' (CP, in this case) of rational beliefs based on precise knowledge of the probability function: in a context in which the total number of lottery tickets is considered to be ultralarge, the probability of one (or a few) tickets is ultrasmall and rounded to zero.

4.4 The epistemology of an infinite lottery

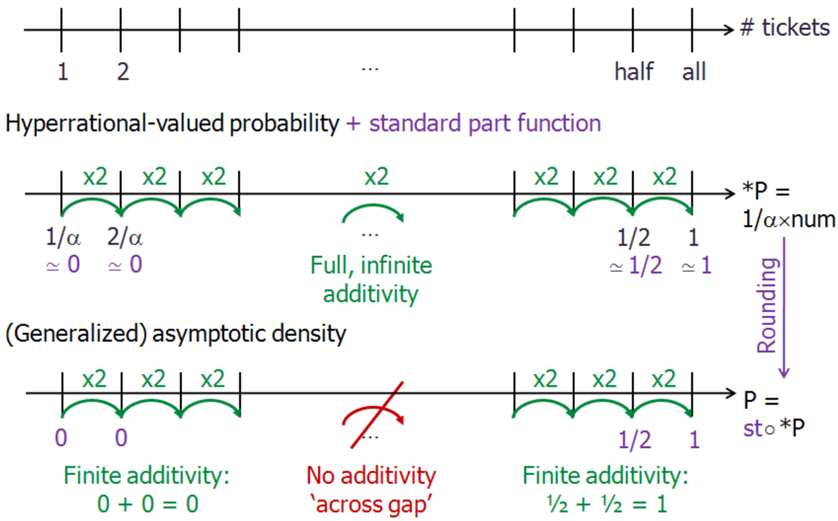
In this final section, we will illustrate how we can combine the two approaches to describe rational beliefs about an infinite lottery. In that case, two rounding processes occur: one in Step 2 and another in Step 3.

Step 1. If we want to determine the winning odds of a specific ticket or set of tickets in a fair, countably infinite lottery, we may express this probability as a hyperrational number, in particular, as a normalized numerosity. This solution is exact (although there is some freedom in the choice of the non-standard model).

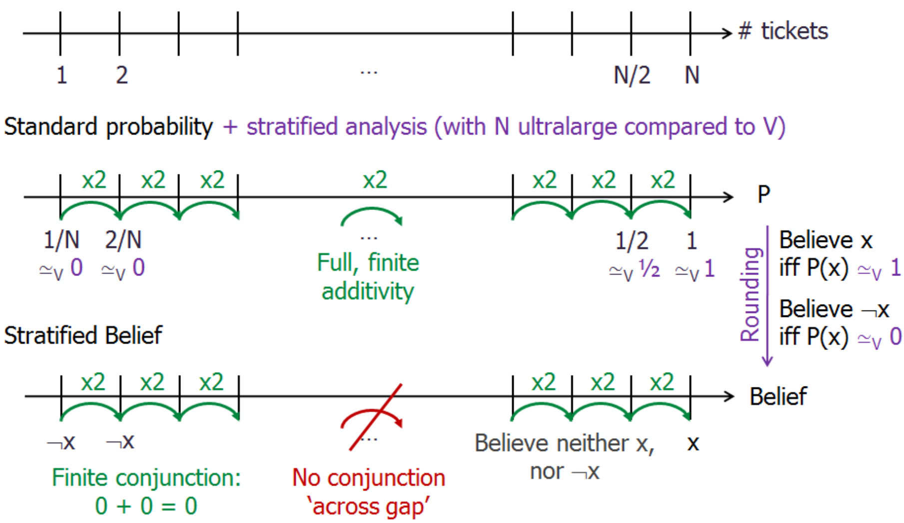
Step 2. If we want to apply our definition of Stratified Belief to a hyperrational probability value, we cannot do so directly, for it works only on real-valued probabilities. Taking the standard part of the outcome, however, results in a real-valued approximation to the probability. In other words, the value is rounded up to an infinitesimal a first time.

Step 3. Subsequently, we can apply the usual definition of Stratified Belief on this already rounded probability value, which results in an additional rounding, this time up to a *relative* infinitesimal.

According to this solution, it is equally rational to believe that your tickets will not win if you own any finite number of tickets in a fair infinite lottery, as it is rational to do so for a finite lottery in which you own 'a few' tickets. Admittedly, the probability of winning is not zero in either case, but beliefs are modeled here as almost certainty, which may be completely false in some (*i.e.* an infinitesimal fraction) of all cases.



(a) Infinite lottery: hyperrational and real probability.



(b) Finite lottery: real probability and Stratified Beliefs.

Figure 4.1: *Illustration of the analogy between (a) an infinite lottery and (b) a large but finite lottery. Before rounding off the (relative) infinitesimals, the probability functions are fully additive. For the rounded values, an appropriately weakened additivity rule applies.*